U.S. DEPARTMENT OF COMMERCE
National Technical Information Service

AD-A035 262

MILITARY CONSTRUCTION ENGINEERING AND DESIGN COST FORECASTS

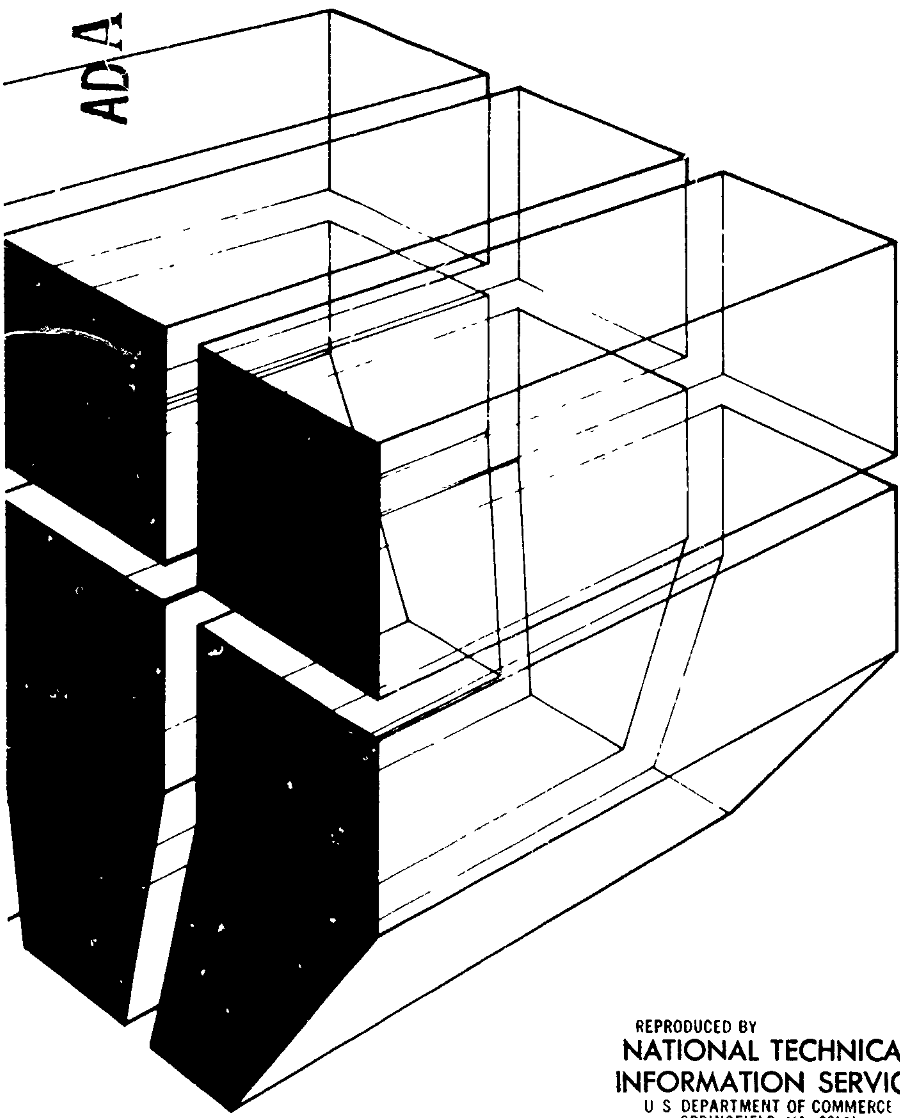
CONSTRUCTION ENGINEERING RESEARCH LABORATORY (ARMY)
CHAMPAIGN, ILLINOIS

JANUARY 1977

construction
engineering
research
laboratory

TECHNICAL REPORT P-77
January 1977

MILITARY CONSTRUCTION ENGINEERING AND DESIGN COST FORECASTS



by
Michael J. O'Connor
Gerald J. Brown
John R. DeCardy

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER CERL-TR-P-77	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) MILITARY CONSTRUCTION ENGINEERING AND DESIGN COST FORECASTS		5. TYPE OF REPORT & PERIOD COVERED FINAL
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Michael J. O'Connor Gerald J. Brown John R. DeCardy		8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS CONSTRUCTION ENGINEERING RESEARCH LABORATORY P.O. Box 4005 Champaign, IL 61820		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 4A762719AT05-03-003
11. CONTROLLING OFFICE NAME AND ADDRESS		12. REPORT DATE January 1977
		13. NUMBER OF PAGES 30
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

Copies are obtainable from National Technical Information Service
Springfield, VA 22151

19 KEY WORDS (Continue on reverse side if necessary and identify by block number)

engineering and design costs
estimated cost of construction
military construction

20 ABSTRACT (Continue on reverse side if necessary and identify by block number)

This report presents a statistical model for forecasting engineering and design (E&D) costs to aid the Directorate of Military Construction in establishing yearly targets for Division/District E&D rates.

Data for nine military Construction Divisions/Districts from fiscal year (FY) 1966 through fiscal year 1975 were analyzed. A statistically significant model for eight Districts was developed and verified by a retrospective test.

Block 20 continued.

E&D costs/rates predicted as a function of the estimated cost of construction for the eight Districts are presented for FY 76 and FY 77.

FOREWORD

This research was conducted for the Directorate of Military Construction, Office of the Chief of Engineers (OCE), under QCR 1.04.001, Project 4A762719AT05, "Initial Investigation in Military Construction Technology," Task 03, "Construction Methods for Military Facilities," Work Unit 003, "MC Engineering and Design Cost Forecasts."

The OCE Technical Monitor was Mr. David Spivey.

The work was performed by the Management Systems Branch, Facility Acquisition and Construction Division (FA), U.S. Army Construction Engineering Research Laboratory (CERL), Champaign, IL. The principal investigator was Mr. Michael J. O'Connor. Dr. Omar E. Rood, Jr. is Chief, Management Systems Branch, and Mr. E. A. Lotz is Chief, FA.

COL J. E. Hays is Commander and Director of CERL and Dr. L. R. Shaffer is Technical Director.

CONTENTS

TABLES

FIGURES

MILITARY CONSTRUCTION ENGINEERING AND DESIGN COST FORECASTS

1 INTRODUCTION

Background

Each year, the Directorate of Military Construction (DMC) of the Office of the Chief of Engineers (OCE) establishes annual targets for each Corps of Engineers Division/District's engineering and design (E&D) costs for military construction. These targets, which are expressed as a percent of the dollar value of construction designed, are currently established by an empirical procedure based on the previous 4 to 5 years of Division/District performance and the estimated cost of construction (ECC) for the next fiscal year.

Objective

The objective of this work is to develop a statistical model for forecasting E&D costs to aid DMC in establishing E&D targets for each of its Divisions and Districts. The model is to be formulated so that easily obtainable data can be used.

Approach

Data available on the variables thought to affect E&D costs were reviewed. A model was formulated based on these variables, and regression analysis was used to fit the model to the data. The fit of the model was checked, and its capability to accurately forecast E&D rates was determined.

2 METHOD OF DATA REVIEW

E&D costs were collected from the OCE Program Review and Analysis: Division and District Performance Data report for the 10-year period FY 66 through FY 75. Appendix A presents the raw data. Prior to FY 71, the estimated cost of construction was defined as 85 percent of the programmed cost, or if the programmed cost was not established, the equivalent cost. In the second quarter of FY 71, this percentage was changed to 90 percent. The raw data were adjusted to reflect this change, as shown in Appendix A.

The following Divisions/Districts were eliminated from consideration because 10 years of data were not available for them:

a. European Division
b. Huntsville Division
c. Kansas City District
d. Norfolk District
e. Los Angeles District

The Mediterranean Division was eliminated because the recent Saudi Arabian Government workload has invalidated its historical data. The following Divisions/Districts had sufficient data for analysis:

a. Baltimore District
b. Fort Worth District
c. New York District
d. Omaha District
e. Alaska District
f. Mobile District
g. Savannah District
h. Sacramento District
i. Pacific Ocean Division

3 MODEL FORMULATION AND ANALYSIS

Based on the initial data review and the detailed formulation procedure shown in Appendix B, the relationship between E&D costs for a particular Division/District, estimated cost of construction, and time was postulated as:

$$D_i = b_{0i} + b_{1i}C + b_{2i}T + b_{3i}TC \qquad \text{[Eq 1]}$$

where D_i = predicted E&D costs for the ith Division/District ($ MIL)
C = estimated cost of construction ($ MIL)
T = time period (FY 66 = 0, FY 67 = 1 . . .)
$TC = T \times C$
$b_{0i}, b_{1i}, b_{2i}, b_{3i}$ = coefficients for the ith Division/District.

The E&D rate (P_i) is defined as $100 \times D_i/C$.

Because of the limited amount of data available for individual Divisions/Districts, a test was made to determine whether the data from some or all of the Districts could be combined because of similar E&D cost performance. (Appendix B details the test procedure.) It was found that the data from five Districts (Alaska, Baltimore, Fort Worth, New York, and Omaha) could be pooled, thereby increasing confidence in the prediction equations. The regres-

Preceding page blank

sion analyses for the remaining Divisions/Districts were based only on their own data.

The postulated model, when regressed against FY 66 through FY 77 data, provided a good fit for eight Districts, explaining 92.4 to 99.8 percent of the variance of the original data (Table 1). The model did not fit the data for Pacific Ocean Division; it explained only 44.8 percent of the variance of the data and hence is considered inadequate. The large variability and erratic behavior in that Division's E&D cost data caused the lack of fit.

4 MODEL RESULTS AND VERIFICATION

Results

Figures 1 through 8 show the FY 76 and FY 77 prediction equations and graphs for the eight Dis-

Table 1

Regression Results

$D = b_0 + b_1C + b_2T$

Div/District	FY Data	b_0	b_1	b_2	Standard Error of Estimate, S($MIL)	Percent of Variance of Original Data Explained by Model R^2
Alaska	1966-1972	−.0215	.0490	.0688	.160	.997
	1966-1973	.3605	.0410	.0048	.153	.998
	1966-1974	.2963	.0424	.0143	.150	.998
	1966-1975	.3483	.0413	.0062	.149	.998
Baltimore	1966-1972	.1826	.0465	.0908	.160	.997
	1966-1973	.1399	.0474	.0906	.153	.998
	1966-1974	.1601	.0479	.0690	.150	.998
	1966-1975	.1835	.0489	.0330	.149	.998
Fort Worth	1966-1972	.6371	.0377	.2537	.160	.997
	1966-1973	.7278	.0363	.2540	.153	.998
	1966-1974	.7178	.0364	.2537	.150	.998
	1966-1975	.8602	.0337	.2733	.149	.998
Mobile	1966-1972	.3067	.0413		.248	.935
	1966-1973	.1165	.0466		.599	.732
	1966-1974	−.1259	.0504		.564	.870
	1966-1975	−.2325	.0520		.532	.924
New York	1966-1972	.6895	.0369	.0545	.160	.997
	1966-1973	.5357	.0398	.0679	.153	.998
	1966-1974	.5545	.0397	.0621	.150	.998
	1966-1975	.5772	.0397	.0531	.149	.998
Omaha	1966-1972	.1257	.0462	.1516	.160	.997
	1966-1973	.1327	.0460	.1550	.153	.998
	1966-1974	.0827	.0465	.1631	.150	.998
	1966-1975	.0675	.0461	.1779	.149	.998
Pacific Ocean Division	1966-1972	2.9405	.0245		1.686	.271
	1966-1973	2.2439	.0300		1.584	.433
	1966-1974	1.7340	.0338		1.540	.499
	1966-1975	1.9040	.0306		1.510	.448
Sacramento	1966-1972	−.0693	.0622		.228	.969
	1966-1973	.0581	.0588		.235	.972
	1966-1974	.1485	.0565		.235	.975
	1966-1975	.2119	.0551		.226	.985
Savannah	1966-1972	1.3901	.0244		.368	.359
	1966-1973	1.3361	.0256		.342	.383
	1966-1974	1.3962	.0246		.317	.562
	1966-1975	.0100	.0459		.497	.925

tricts for which the model provided a good fit. The prediction limits indicate the accuracy of the prediction. The following example illustrates how they are interpreted. For Baltimore District (Figure 2), if the estimator assumes an estimated cost of construction of $160 million in FY 76, he/she can be 95 percent confident that the actual E&D cost will be between $8.0 and $8.8 million. Equivalently, he/she can be 95 percent confident that the actual E&D rate will be between 5.0 and 5.5 percent.

The graphs have been plotted over a reasonable range of estimated construction costs for each District, since excessive extrapolation is risky. For example, New York District, which has been designing an average of $40 to $45 million of construction per year, would not be expected to perform in the same manner if it were suddenly tasked to design $200 million of construction per year.

Verification

The validity of a forecasting model is determined by its capability to predict future performance. A systematic approach to verifying the model is to conduct a retrospective test. Since lack of the current year's data when the next year's E&D limits must be established may require using the model to predict performance 2 years in the future, the model was tested for predicting both 1 and 2 years ahead. The retrospective test was conducted by comparing the actual FY 74 E&D costs to the FY 74 costs predicted by the model developed from FY 66 through FY 72 data (a 2-year-ahead prediction), and to the FY 74 costs predicted by the model developed from FY 66 through FY 73 data (a 1-year-ahead prediction). Analogous comparisons were also made between actual FY 75 E&D costs and the 1- and 2-year-ahead predictions.

Table 2 presents the results of the retrospective test. In all cases but one, the actual E&D cost was well within the 95 percent confidence limits. Savannah District's actual E&D cost for FY 75 was above the upper prediction limit. This result is within expectations because 95 percent confidence means that approximately 95 out of 100 actual E&D costs will fall within the upper and lower prediction limits.

5 MODEL USE AND MAINTENANCE

Model Use

The model can be used to predict the performance of a particular Division/District based on its past

Table 2

Retrospective Test Results

			1-Year-Ahead Prediction			2-Year-Ahead Prediction		
Div/District	FY	Actual E&D Cost	Predicted E&D Cost	Lower Prediction Limit	Upper Prediction Limit	Predicted E&D Cost	Lower Prediction Level	Upper Prediction Level
Alaska	1974	1 681	1.601	1.064	2.138	1.965	1.035	2.896
	1975	1.444	1.551	1.139	1.963	1.493	.917	2.070
Baltimore	1974	7 932	8.089	7.648	8.531	8.008	7.516	8.500
	1975	6.594	6.906	6.460	7.352	7.017	6.462	7.572
Fort Worth	1974	8.055	8.037	7.563	8.511	8.145	7.460	8.830
	1975	10.141	10.467	9.882	11.051	10.448	9.663	11.232
Mobile	1974	6.591	6.174	3.881	8.467	5.668	4.642	6.695
	1975	7.495	7.236	5.384	9.087	6.928	4.283	9.574
New York	1974	2.735	2.822	2.422	3.222	2.739	2.164	3.315
	1975	2.618	2.783	2.405	3.161	2.822	2.394	3 250
Omaha	1974	12.386	12.050	11.645	12.455	12.077	11.529	12.625
	1975	11 422	11.314	10.910	11.719	11.185	10.740	11.630
Sacramento	1974	5 068	5.370	4.633	6.107	5.558	4.703	6.414
	1975	6 721	6 924	6.132	7.716	7.110	6 174	8.046
Savannah	1974	3.970	4.029	2.392	5.666	3.959	2 052	5.866
	1975	8 266	5 567	3.779	7 355	5.674	2.101	9.248

performance. DMC must exercise its managerial control in establishing E&D targets which may or may not be equal to the predicted rate. Since District performance is influenced by the established target, DMC may predict the expected E&D rate and then establish the target either below or above this value to restrict or expand the E&D effort, as desired. For example, if restrictions are deemed necessary for a particular District based on past performance, the DMC may wish to establish the target at the lower 95 percent confidence limit.

The prediction equations given for each of the Districts in Table 1 (FY 66 through FY 75 data) can be used to calculate an estimated E&D cost by substituting the estimated cost of construction and the appropriate time period (FY 76 = 10, FY 77 = 11, etc.). The E&D rate is calculated by dividing the E&D cost by the estimated cost of construction. In Figures 1 through 8, the value of time is already accounted for and combined with the constant term; therefore, only an estimated cost of construction is required. For example, Baltimore's prediction equation is:

$$D = .1835 + .0489\,(C) + .0330\,(T)$$

Assuming estimated construction costs for FY 76 at \$160 million,

$$D = .1835 + .0489\,(160)^{*} + .0330\,(10).^{**}$$

Therefore,

$$D = \$8.337 \text{ million}$$

and

$$P = \frac{\$8.337}{\$160} \times 100 = 5.2 \text{ percent.}$$

Based on this value of expected actual performance, the established limit may be set according to the current management policy for the District.

Model Maintenance

The stability of the model parameters, the good fit of the model with the data, and the accuracy with which the model is predicting actual E&D costs should be checked periodically. Table 1 shows the parameter changes as additional data are added, beginning with the data through FY 72. The coefficient for the estimated cost of construction remains fairly stable; however, in some Districts, such as Alaska, the other parameters change as additional data are added. This indicates that although the basic model may not be changing, the effects of the variables—particularly time—on E&D costs may be, and caution should be exercised in predicting several years in advance. New parameters should be determined when each year's data are added.

*160 is used because the model is based on millions of dollars.
**FY 76 = 10.

The goodness of fit of the model can be checked by examining the estimated standard error of the estimate, S, and the percent of variance in the original data explained by the model, R^2 (Table 1). An increasing value of S or a decreasing value of R^2 would indicate that the basic model itself may be changing and steps should be taken to determine where changes are occurring. Stable or decreasing trends in S indicate a stable model. The high values for R^2 indicate the model is describing the data very well.

Finally, the accuracy of the model as a predictor should be checked annually. Actual values which continually fall outside the confidence limits may indicate changes in the model itself.

6 CONCLUSIONS AND RECOMMENDATIONS

The model presented in this report is a valid technique for forecasting future E&D costs. Predictions were presented for FY 76/FY 77 for the eight Military Construction Divisions/Districts which had sufficient data for analysis.

It is recommended that the DMC use this model as an aid in establishing E&D targets. Although the model can be used to predict costs 1 or 2 years in advance, the best results are obtained for predictions made 1 year ahead. Thus, the model should be updated with the past year's data before determining the new year's targets.

Because the E&D costs predicted by the model are based on past Division/District performance and conditions, significant changes in performance or conditions could invalidate the historical data on which the model is based. Therefore, the DMC must consider changes in operating conditions or management policy within the Division/District when establishing E&D targets.

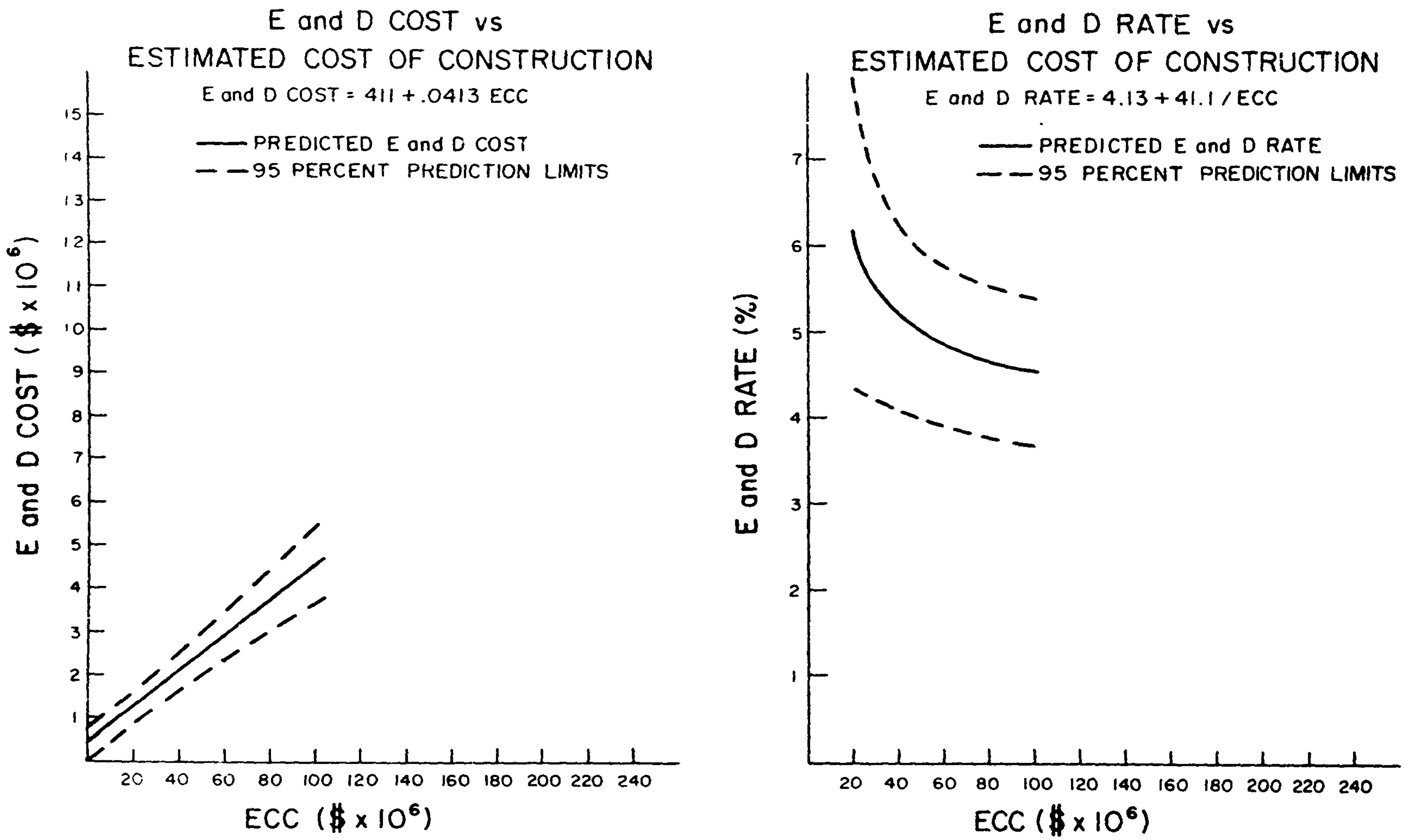


Figure 1a. Prediction of E&D costs/rates—Alaska—FY 76.

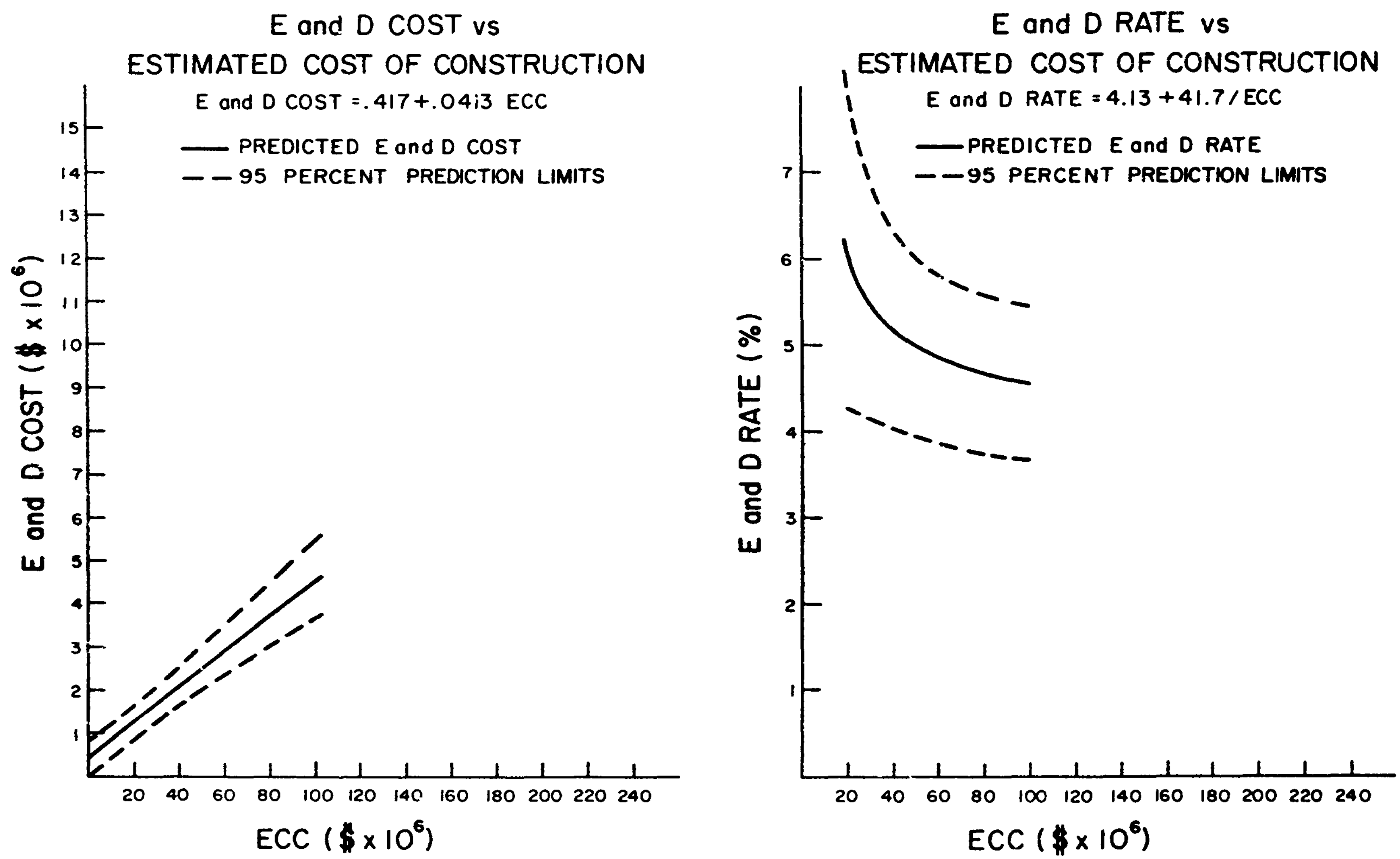


Figure 1b. Prediction of E&D costs/rates—Alaska—FY 77.

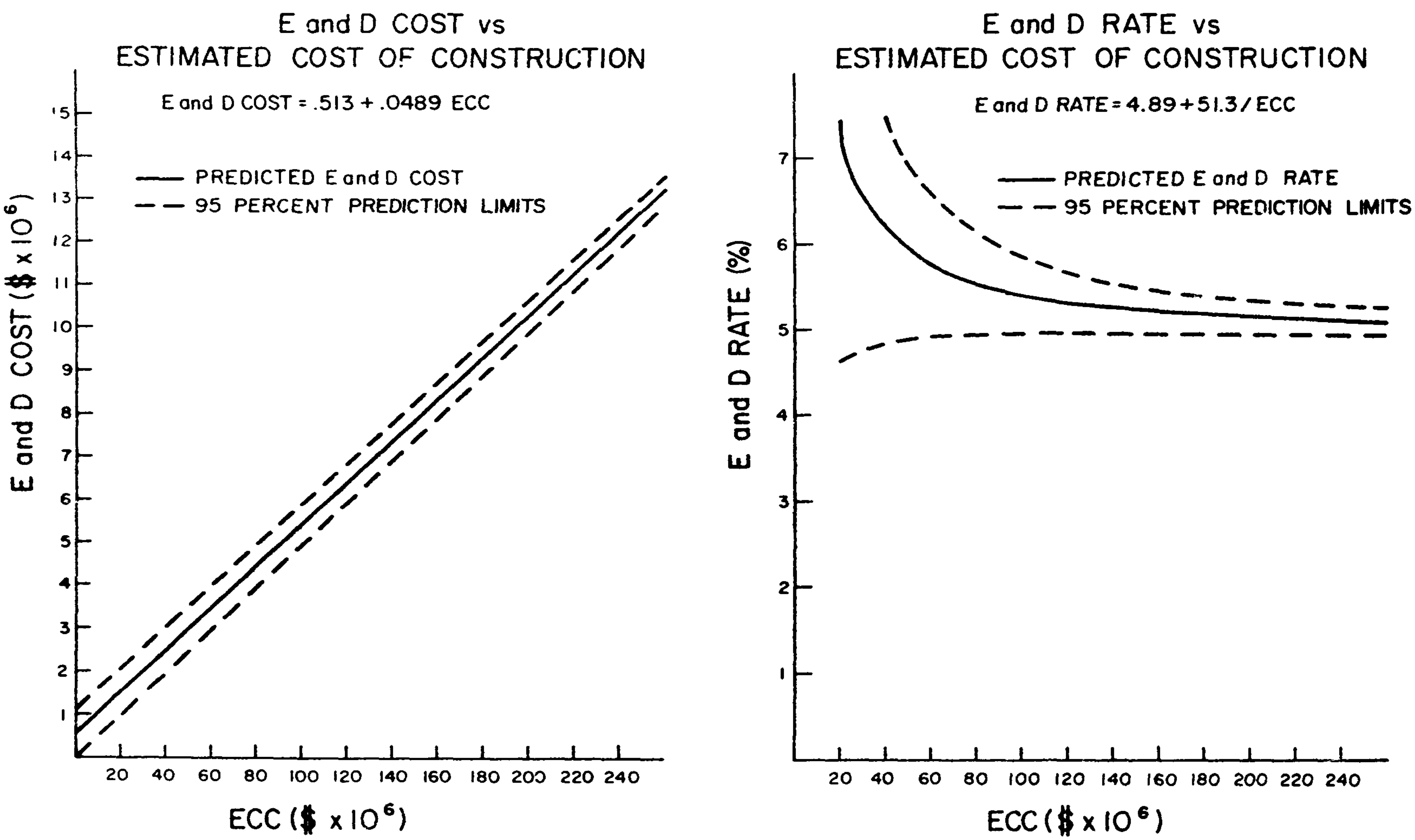


Figure 2a. Prediction of E&D costs/rates—Baltimore—FY 76.

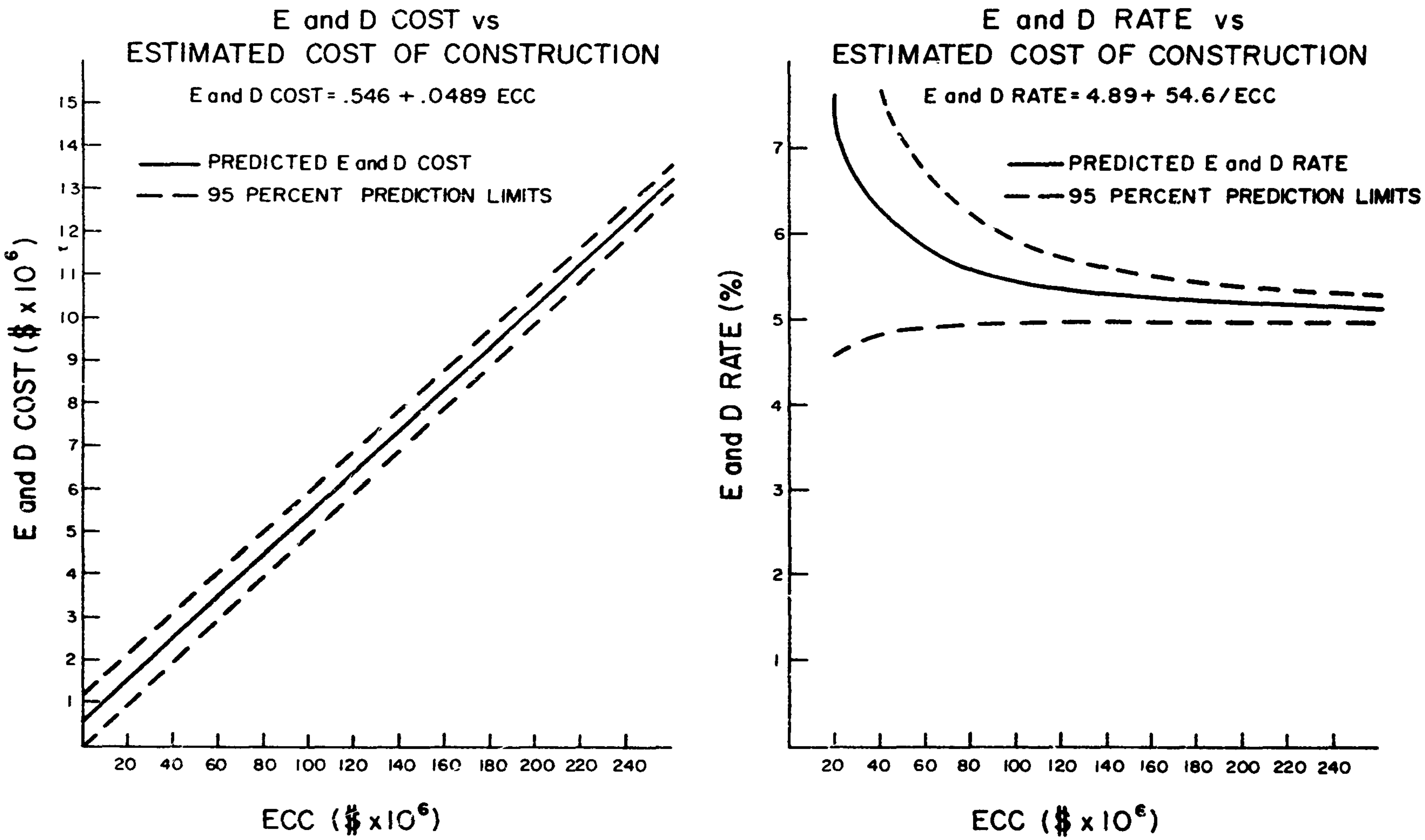


Figure 2b. Prediction of E&D costs/rates—Baltimore—FY 77.

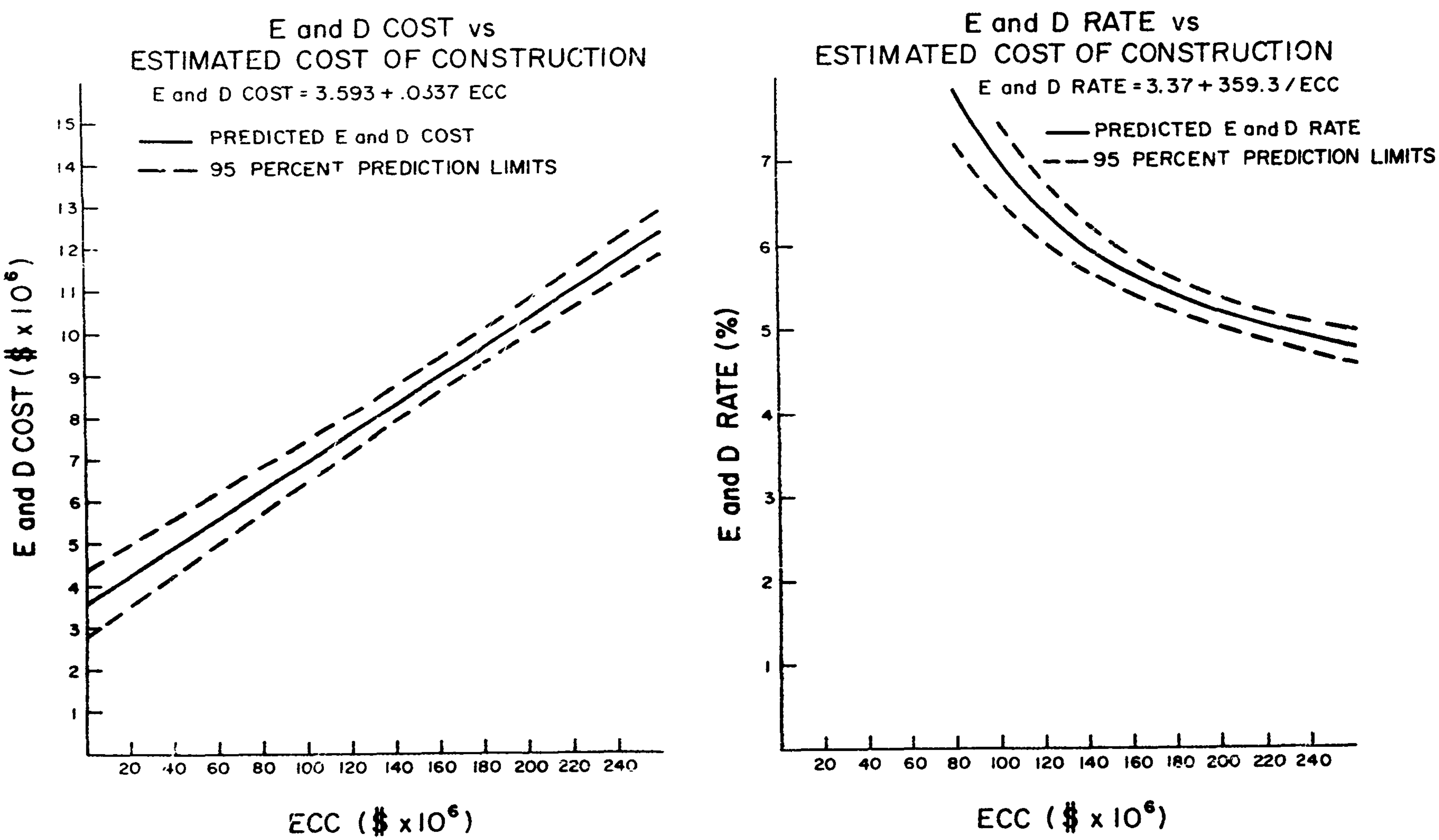


Figure 3a. Prediction of E&D costs/rates—Fort Worth—FY 76.

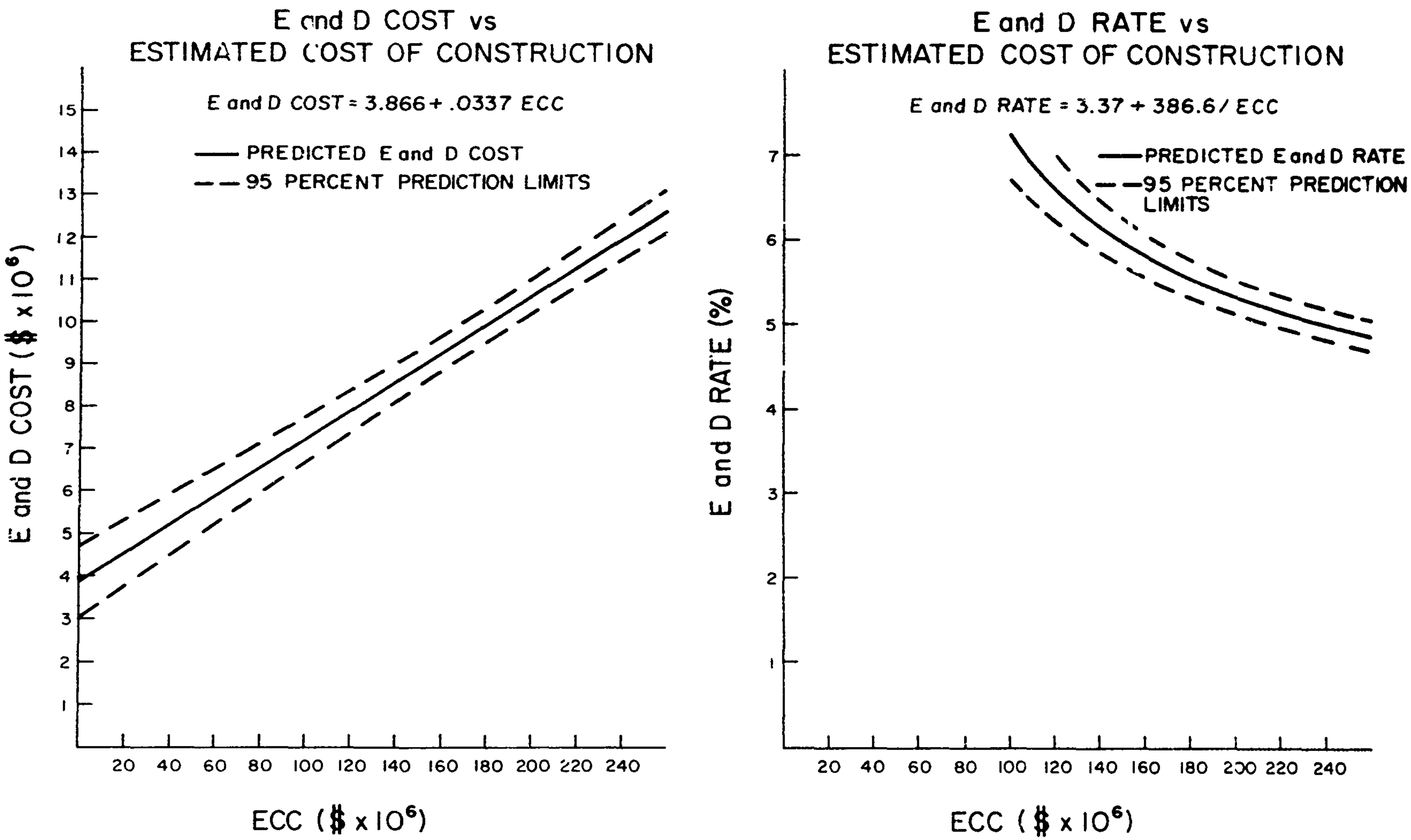


Figure 3b. Prediction of E&D costs/rates—Fort Worth—FY 77.

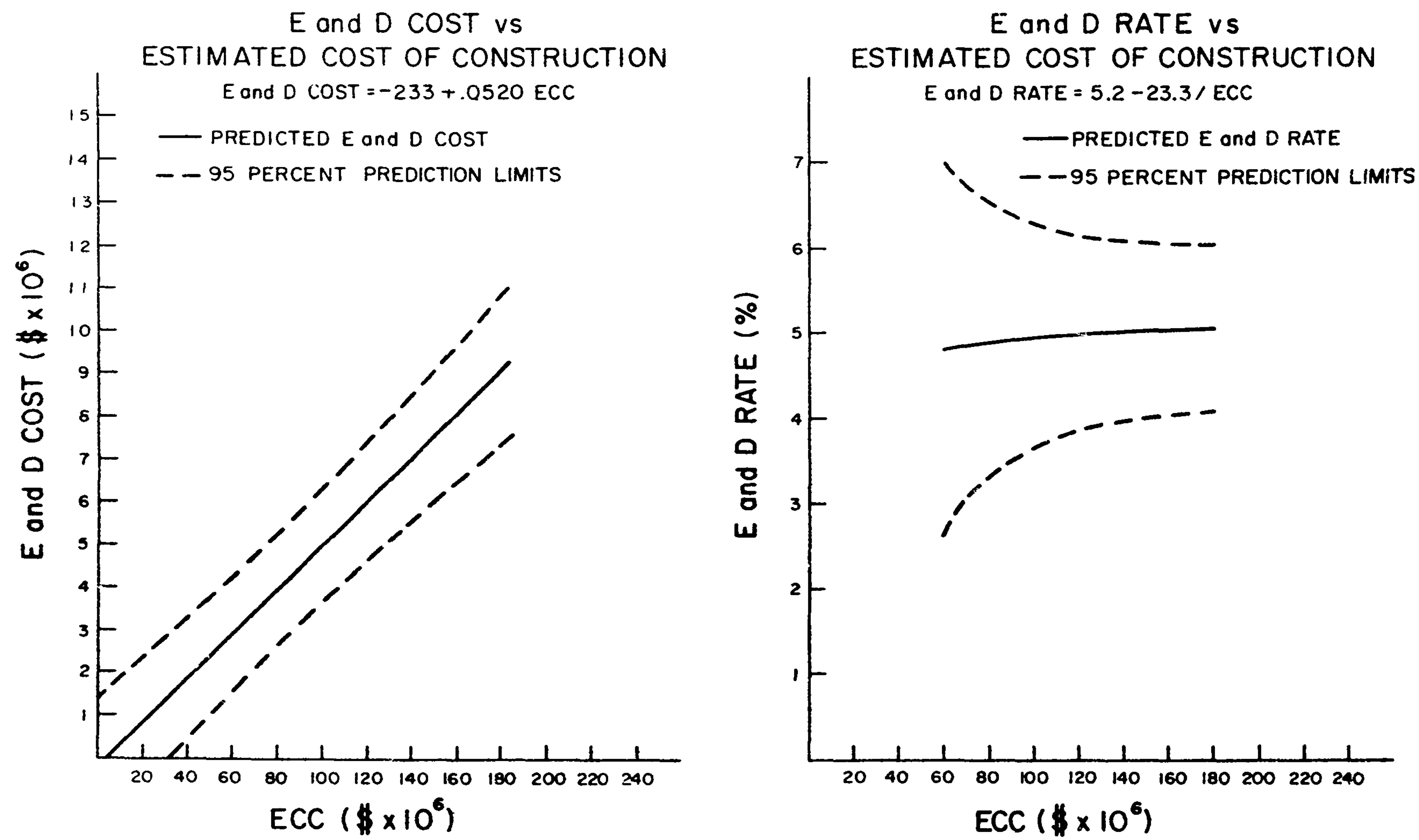


Figure 4. Prediction of E&D costs/rates—Mobile—FY 76/FY 77.

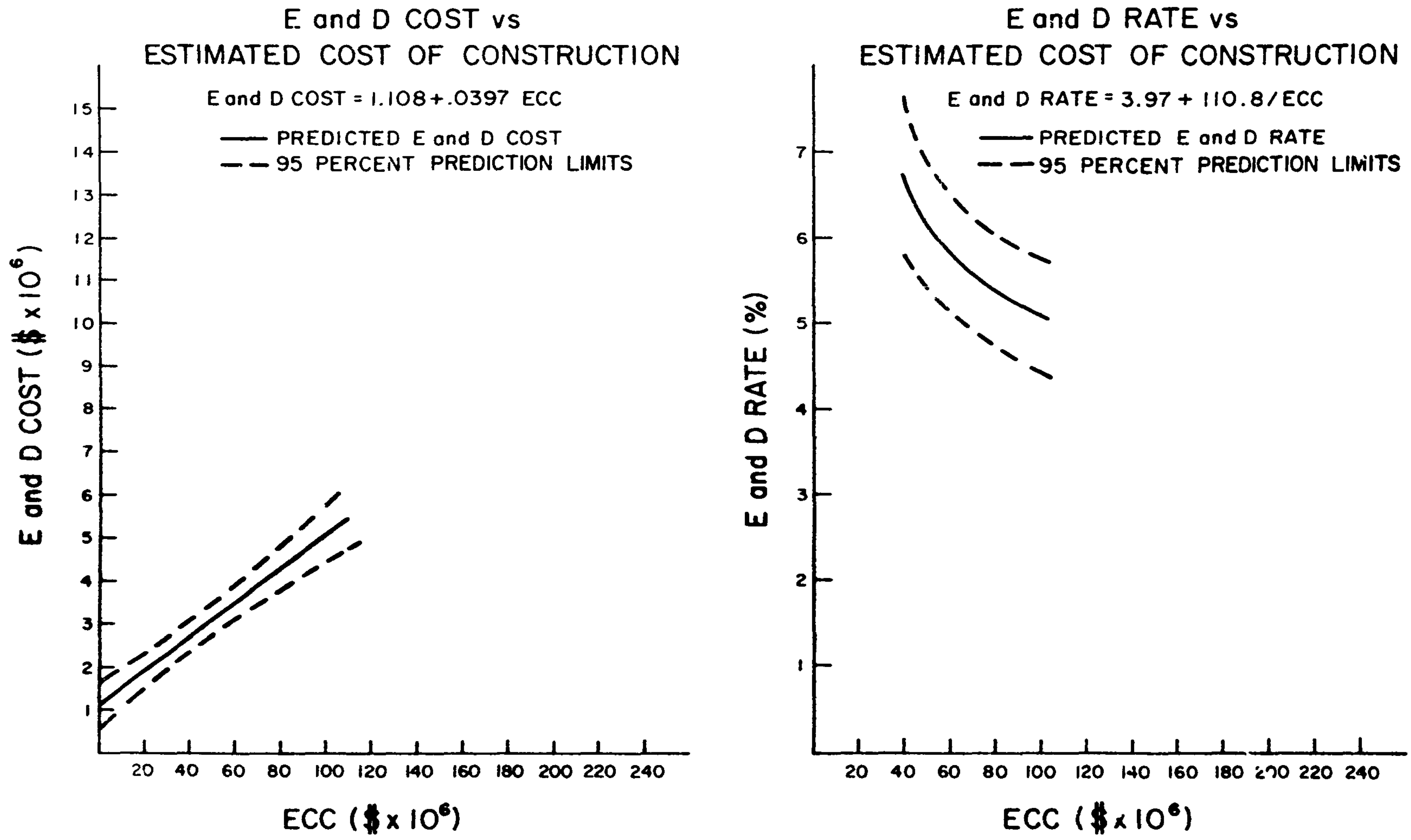


Figure 5a. Prediction of E&D costs/rates—New York—FY 76.

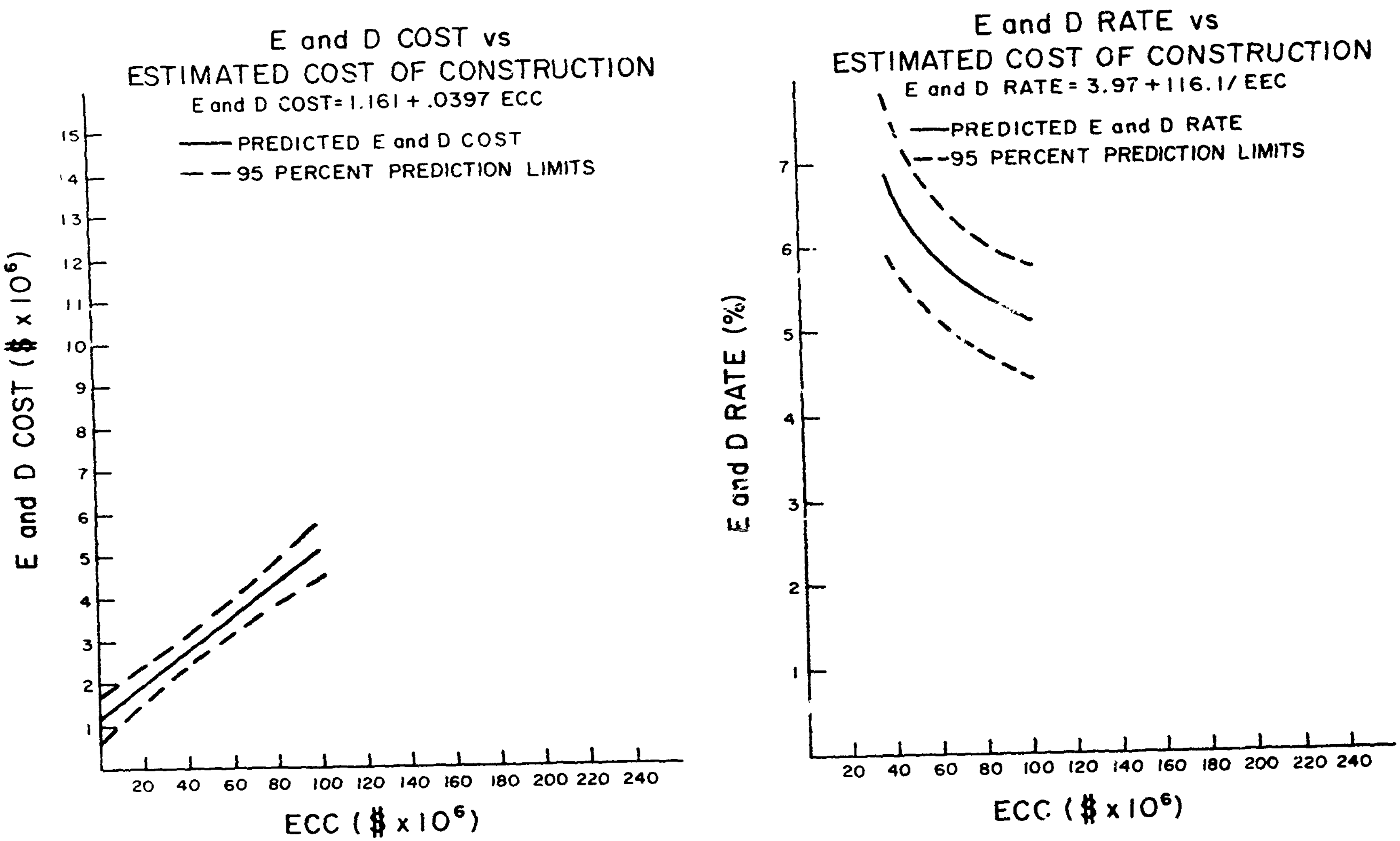


Figure 5b. Prediction of E&D costs/rates—New York—FY 77.

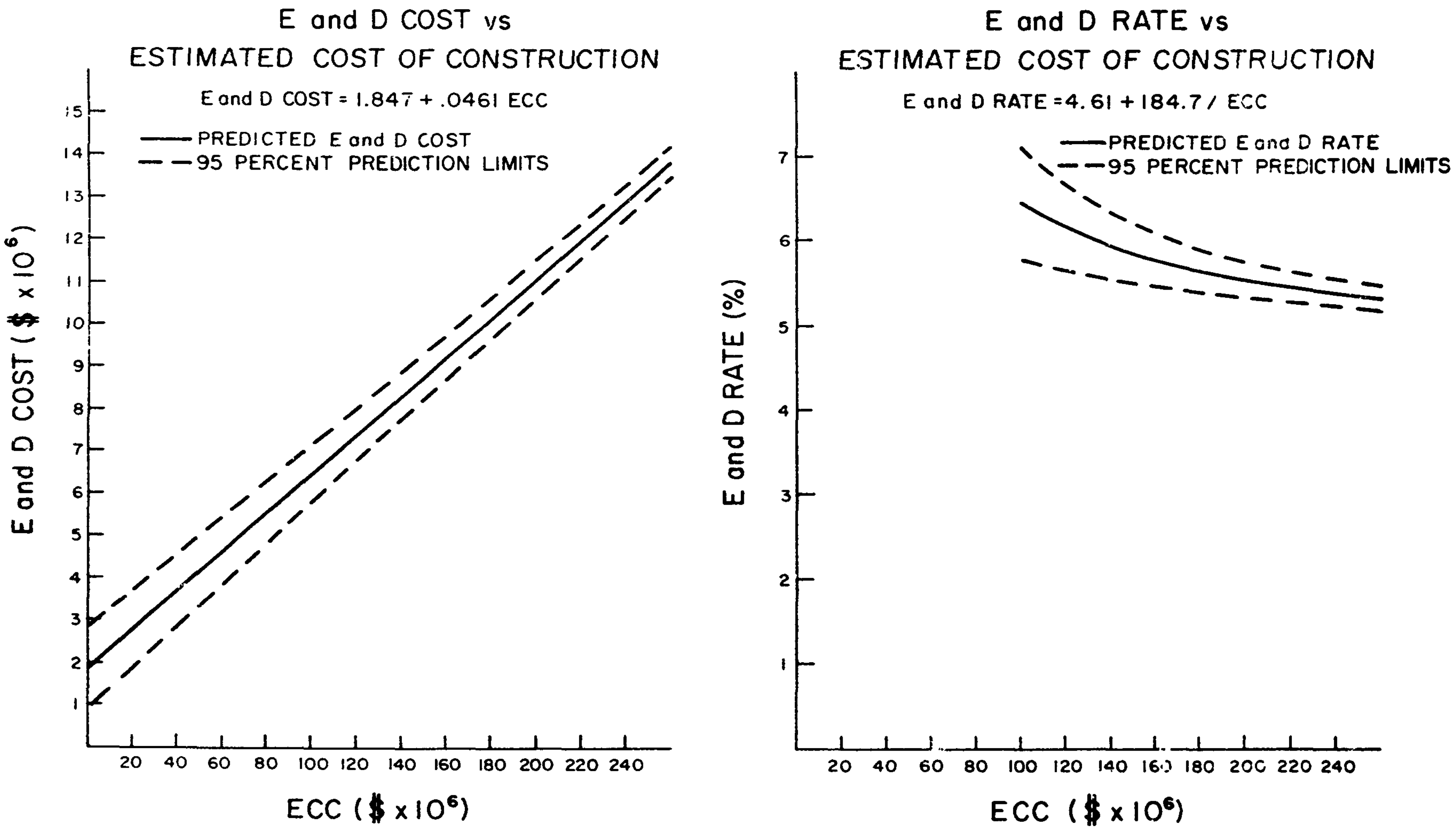


Figure 6a. Prediction of E&D costs/rates—Omaha—FY 76.

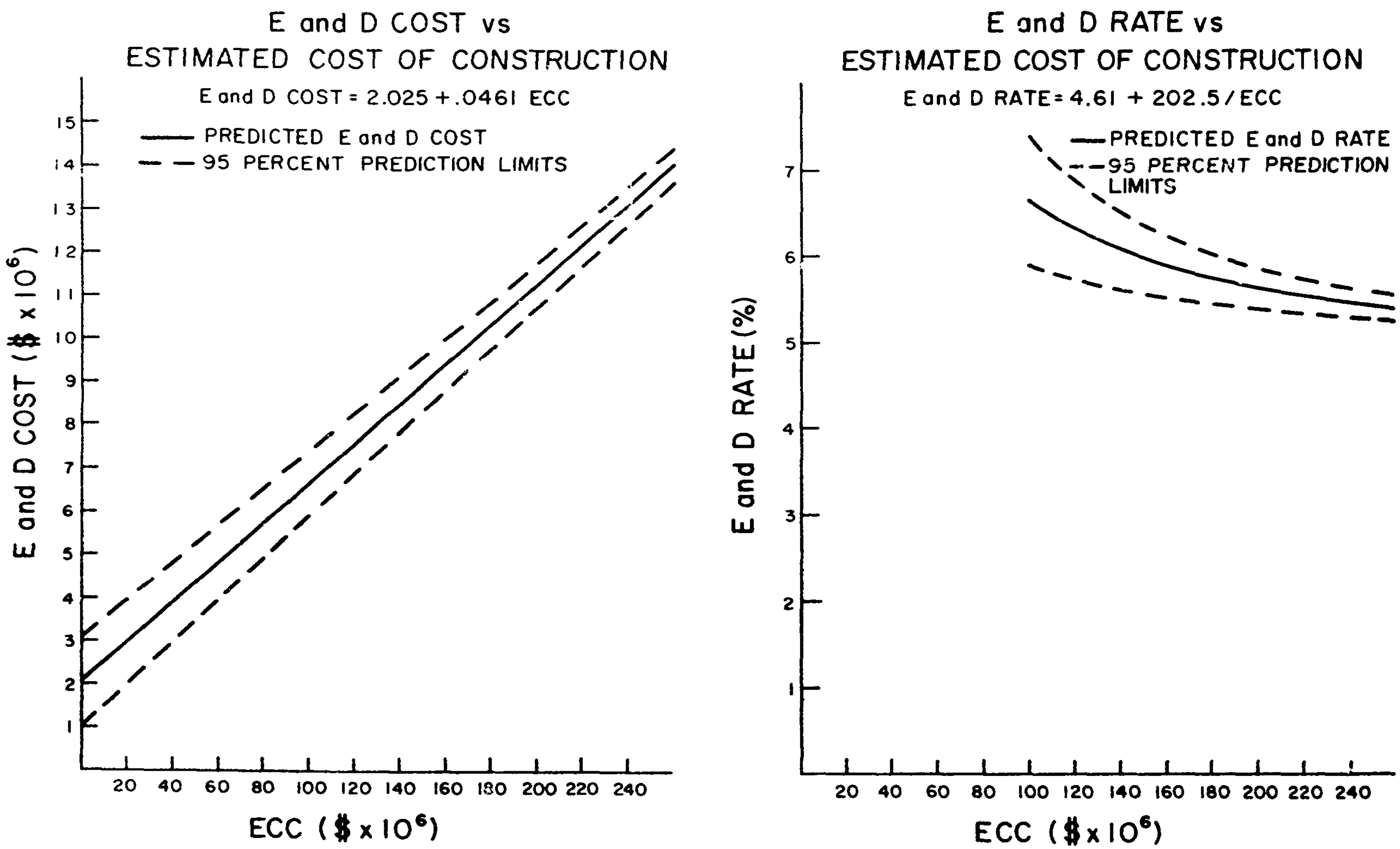


Figure 6b. Prediction of E&D costs/rates—Omaha—FY 77.

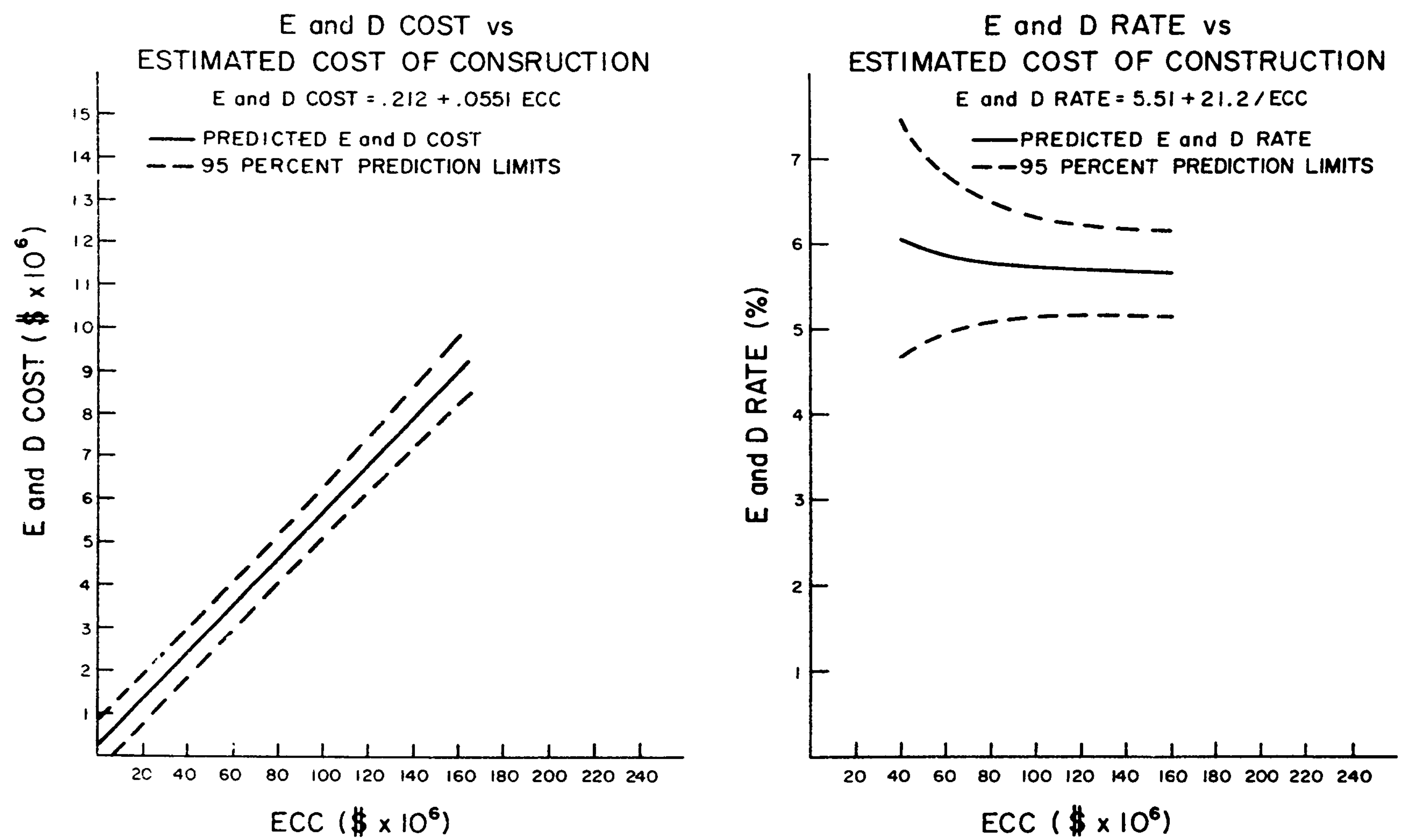


Figure 7. Prediction of E&D costs/rates—Sacramento—FY 76/FY 77.

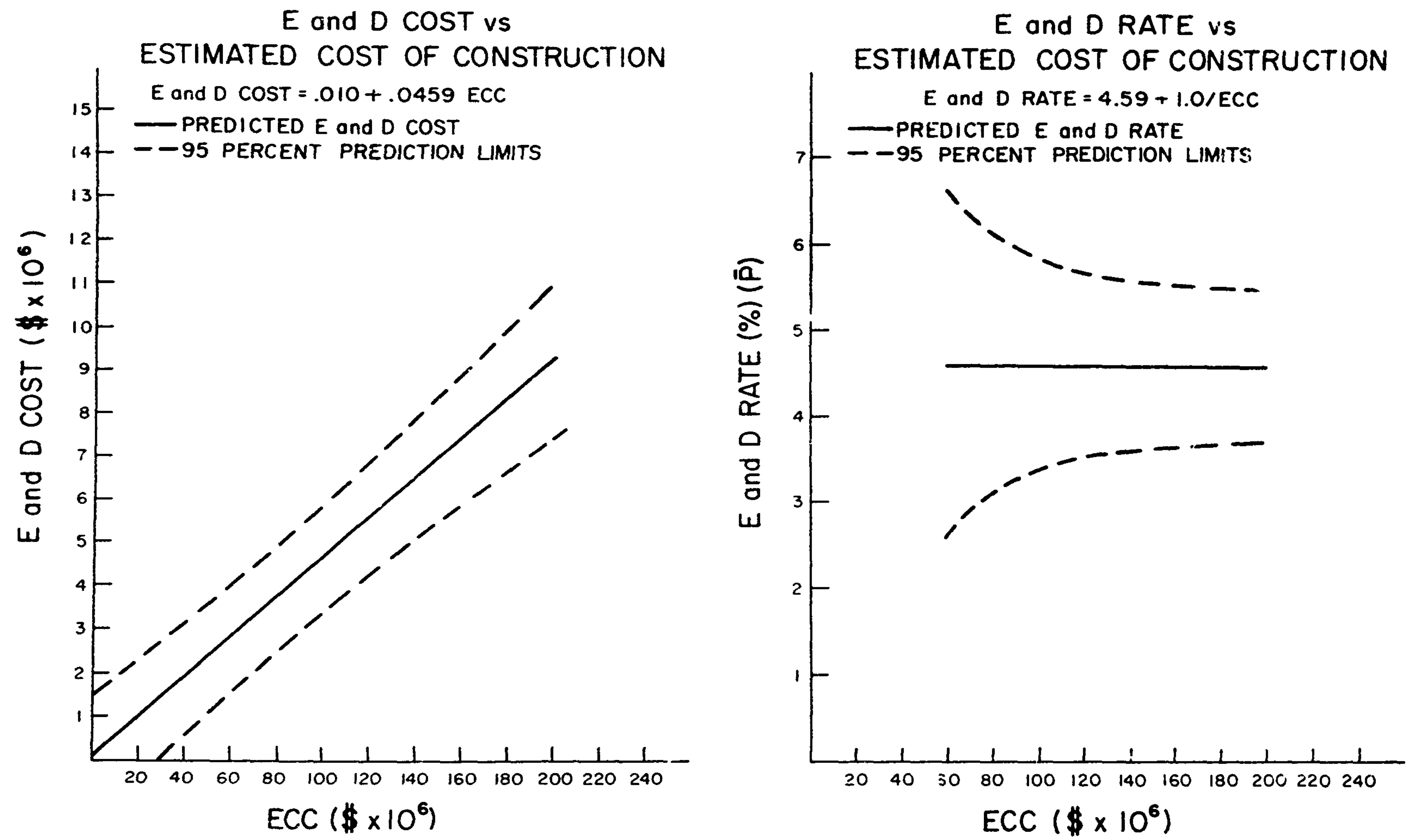


Figure 8. Prediction of E&D costs/rates—Savannah—FY 76/FY 77.

APPENDIX A:

DATA FOR FY 66 THROUGH FY 75

ALASKA DISTRICT

FY	ECC ($ Mil)	Adj ECC ($ Mil)	E&D Cost ($ Mil)	E&D Rate (%)
1966	36.509	38.656	1.740	4.501
1967	41.840	44.301	2.120	4.785
1968	38 756	41.036	2.284	5.566
1969	16.526	17.498	1.157	6.612
1970	15.326	16.238	1.201	7.396
1971	11.550	11.550	.867	7.506
1972	8.156	8.156	.612	7.504
1973	21.025	21.025	1.170	5.565
1974	29.340	29.340	1.681	5.729
1975	26.589	26.589	1.444	5.431

FORT WORTH DISTRICT

FY	ECC ($ Mil)	Adj ECC ($ Mil)	E&D Cost ($ Mil)	E&D Rate (%)
1966	63.034	66.742	3.285	4.922
1967	56.627	59.958	2.960	4.937
1968	52.366	55.446	3.168	5.714
1969	74.778	79.176	4.435	5.601
1970	53.133	56.258	3.969	7.055
1971	75.350	75.350	4.598	6.102
1972	104.900	104.900	6.128	5.842
1973	122.985	122.985	6.921	5.628
1974	145.416	145.426	8.055	5.539
1975	204.842	204.842	10.141	4.951

BALTIMORE DISTRICT

FY	ECC ($ Mil)	Adj ECC ($ Mil)	E&D Cost ($ Mil)	E&D Rate (%)
1966	45.126	47.780	2.439	5.105
1967	29.759	31.509	1.701	5.398
1968	40.307	42.678	2.281	5.345
1969	33.714	35.697	2.228	6.241
1970	44 262	46.865	2.654	5.663
1971	130 216	130.216	6.778	5.205
1972	186 851	186 851	9.376	5.018
1973	214 391	214.391	11 053	5.156
1974	152 560	152.560	7 932	5.199
1975	128.000	128.000	6.594	5.152

HUNTSVILLE DIVISION

FY	ECC ($ Mil)	Adj ECC ($ Mil)	E&D Cost ($ Mil)	E&D Rate (%)
1966				
1967				
1968				
1969				
1970	204.532	216.563	8.184	3.779
1971	241.000	241.000	9.773	4.055
1972	216.700	216.700	9.369	4.323
1973	91.836	91.836	4.875	5.308
1974	48.527	48.527	2.426	4.999
1975	180.712	180.712	11.041	6.110

EUROPEAN DIVISION

FY	ECC ($ Mil)	Adj ECC ($ Mil)	E&D Cost ($ Mil)	E&D Rate (%)
1966	16.049	16.993	1.029	6.055
1967	25.176	26.657	1.588	5.957
1968	27 652	29 278	1.826	6.237
1969	10.830	11 467	769	6.706
1970	12.853	18.080	1.086	8.303
1971				
1972				
1973				
1974				
1975	137 713	137.713	11.041	8.017

KANSAS CITY DISTRICT

FY	ECC ($ Mil)	Adj ECC ($ Mil)	E&D Cost ($ Mil)	E&D Rate (%)
1966	22.335	23.649	931	3.937
1967	27.840	29.478	1.300	4.410
1968	62.094	65.746	1.939	2.949
1969	50.116	53.064	2.169	4.088
1970	20.920	22.151	1.276	5.761
1971				
1972				
1973				
1974				
1975	5.843	5.843	315	5.391

LOS ANGELES DISTRICT

FY	ECC ($ Mil)	Adj ECC ($ Mil)	E&D Cost ($ Mil)	E&D Rate (%)
1966	33.938	35.934	1.910	5.315
1967	40.013	42.367	2.284	5.391
1968	44.800	47.435	2.652	5.591
1969	57.364	60.738	4.039	6.650
1970	33.373	35.336	2.336	6.611
1971	3.257	3.257	.318	9.764
1972	.658	.658	.110	16.717
1973				
1974				
1975				

NEW YORK DISTRICT

FY	ECC ($ Mil)	Adj ECC ($ Mil)	E&D Cost ($ Mil)	E&D Rate (%)
1966	52.190	55.260	2.723	4.928
1967	47.350	50 135	2.549	5.084
1968	34.493	36.522	2.035	5.572
1969	33.334	35.295	2.259	6.400
1970	30.393	32.181	2.382	7.402
1971	27.104	27.104	1.735	6.401
1972	44.376	44.376	2.638	5.945
1973	60.947	60.947	3.470	5.693
1974	43.805	43.805	2.735	6.244
1975	42.084	42.084	2.618	6.221

MEDITERRANEAN DIVISION

FY	ECC ($ Mil)	Adj ECC ($ Mil)	E&D Cost ($ Mil)	E&D Rate (%)
1966	42.518	45.019	1.727	3.836
1967	81 305	86.087	3.490	4.054
1968	49.106	51.994	2.076	3.993
1969	36.544	35.694	1.248	3.225
1970	20.820	22.045	.907	4.114
1971	13.844	18.844	.988	7.187
1972	27.445	27.445	1.871	6.817
1973	20.695	20.695	1.646	7.954
1974	96.356	96.356	3.165	3.285
1975	1259 372	1259 372	9.505	.755

NORFOLK DISTRICT

FY	ECC ($ Mil)	Adj ECC ($ Mil)	E&D Cost ($ Mil)	E&D Rate (%)
1966	31.611	33.470	1.739	5.196
1967	42.120	44.597	2.361	5.294
1968	53.869	57.038	2.968	5.204
1969	28.139	29.794	1.993	6.689
1970	29.068	30.778	2.306	7.492
1971				
1972				
1973				
1974	69.945	69.945	2.730	3.903
1975	76.553	76.553	3.522	4.601

MOBILE DISTRICT

FY	ECC ($ Mil)	Adj ECC ($ Mil)	E&D Cost ($ Mil)	E&D Rate (%)
1966	72 269	76 520	3.206	4.190
1967	79 823	84.518	4.219	4.992
1968	49 803	52.732	2.370	4.494
1969	30 889	32 706	1 821	5 568
1970	63.821	67.575	2.976	4.404
1971	71.261	71 261	3.162	4.437
1972	95 799	95.799	4.241	4 427
1973	79.929	79.929	5.087	6.364
1974	129 964	129 964	6.591	5.071
1975	146.148	146 148	7.495	5.128

OMAHA DISTRICT

FY	ECC ($ Mil)	Adj ECC ($ Mil)	E&D Cost ($ Mil)	E&D Rate (%)
1966	37.368	39.778	2.085	5.242
1967	29.604	31.345	1.572	5.015
1968	44.580	47.202	2.560	5 423
1969	28 811	29.976	2.075	6 922
1970	72.893	77.181	4.211	5.456
1971	123.120	123.120	6.697	5.439
1972	167 840	167 840	8.747	5.212
1973	228 175	228 175	11.694	5.125
1974	232.203	232.203	12.386	5.334
1975	210.027	210.027	11.422	5.438

PACIFIC OCEAN DIVISION

FY	ECC ($ Mil)	Adj ECC ($ Mil)	E&D Cost ($ Mil)	E&D Rate (%)
1966	143.306	151.735	5.275	3.476
1967	139.185	147.372	5.423	3.680
1968	68.728	93.947	4.944	5.263
1969	135.920	143.915	9.132	6.345
1970	73.209	77.515	6.490	8.373
1971	66.213	66.213	3.871	5.846
1972	74.332	74.332	3.970	5.341
1973	40.445	40.445	2.743	6.782
1974	56.003	56.003	2.527	4.512
1975	133.867	133.867	4.844	3.619

SAVANNAH DISTRICT

FY	ECC ($ Mil)	Adj ECC ($ Mil)	E&D Cost ($ Mil)	E&D Rate (%)
1966	67.693	71.675	2.723	3.799
1967	67.615	71.592	2.886	4.031
1968	55.518	58.784	2.745	4.670
1969	58.247	61.673	3.005	4.872
1970	59.922	63.447	3.489	5.499
1971	41.497	41.497	2.201	5.304
1972	66.512	66.512	3.2[illegible]5	4.954
1973	67.684	67.684	3.204	4.734
1974	105.320	105.320	3.970	3.769
1975	169.654	169.654	8.266	4.872

SACRAMENTO DISTRICT

FY	ECC ($ Mil)	Adj ECC ($ Mil)	E&D Cost ($ Mil)	E&D Rate (%)
1966	25.684	27.195	1.609	5.917
1967	23.844	25.247	1.427	5.652
1968	46.387	49.115	2.637	5.369
1969	33.090	35.036	2.248	6.416
1970	34.985	37.043	2.453	6.622
1971	71.446	71.446	4.598	6.436
1972	68.049	68.049	4.035	5.930
1973	81.456	81.456	4.643	5.700
1974	90.403	90.403	5.068	5.606
1975	120.013	120.013	6.721	5.600

APPENDIX B:

MATHEMATICAL DEVELOPMENT

Model Formulation

Two variables were available for use in estimating E&D costs, D, for a particular Division/District: (1) estimated cost of construction, C, and (2) time, T. The first step was to determine the relationship between D, C, and T. A correlation matrix (Table B1) was computed from the adjusted data.

Table B1

Correlation of Dependent/Independent Variables

	D	C	T
D	1.000	.969	.482
C	.969	1.000	.437
T	.482	.437	1.000

The initial Division/District model at a base time was formulated as a linear relationship between D and C:

$$D = a_o + a_1 C \qquad \text{[Eq B1]}$$

The E&D costs equal some fixed costs, a_o, plus some fixed percentage, a_1, of the estimated cost of construction (Figure B1).

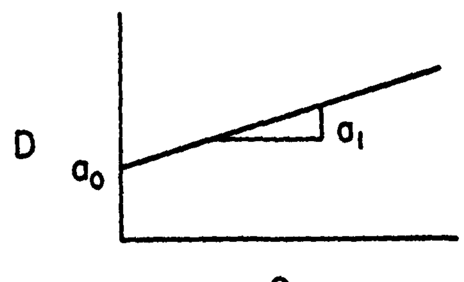


Figure B1. E&D costs vs estimated cost of construction.

The average E&D rate ($\overline{P}$) is defined as the E&D cost divided by the estimated cost of construction designed times 100:

$$\overline{P} = \frac{100D}{C} \qquad \text{[Eq B2]}$$

The model can then be interpreted in terms of average percentages by dividing Eq B1 by C:

$$\overline{P} = 100\, a_1 + 100\, a_o/C \qquad \text{[Eq B3]}$$

Figure B2 is a graphical interpretation of Eq B3.

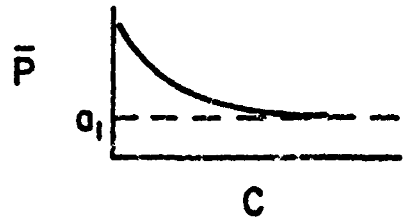


Figure B2. E&D rate vs estimated cost of construction.

The average percent E&D decreases as the estimated cost of construction increases. This phenomenon, which is known as "economy of scale," is due to the amortization of the fixed charge over a larger workload base. If there is no economy of scale within a Division/District, the general model would automatically collapse to Eq B4 and B5 and Figure B3.

$$D = a_1 C \qquad \text{[Eq B4]}$$

$$\overline{P} = 100\, a_1 \qquad \text{[Eq B5]}$$

The next variable to be considered was time. Costs of both design and construction increase over time, probably at different rates. Some of these costs are:

a. Materials
b. Labor
c. Effects of Environmental Impact Assessments/Statements
d. Workload levels
e. Manpower levels.

To account for changes in D and C across time the total change in each was modeled by uniform annual rates compounded over time. The advantage of this method is that it does not require the imposition of indices that may not completely correspond to the inflationary pressures exerted on the type of design and construction that the Corps performs. These uniform annual rates reflect not only inflationary effects on D and C, but also noninflationary effects such as changes in workload or manpower levels.

The uniform annual rates of change in design and construction costs can be represented as i_d and i_c, respectively, where

$$D_t = (1 + i_d)^t D \quad t = 0, 1, \ldots, n-1 \qquad \text{[Eq B6]}$$

$$C_t = (1 + i_c)^t C \quad t = 0, 1, \ldots, n-1 \qquad \text{[Eq B7]}$$

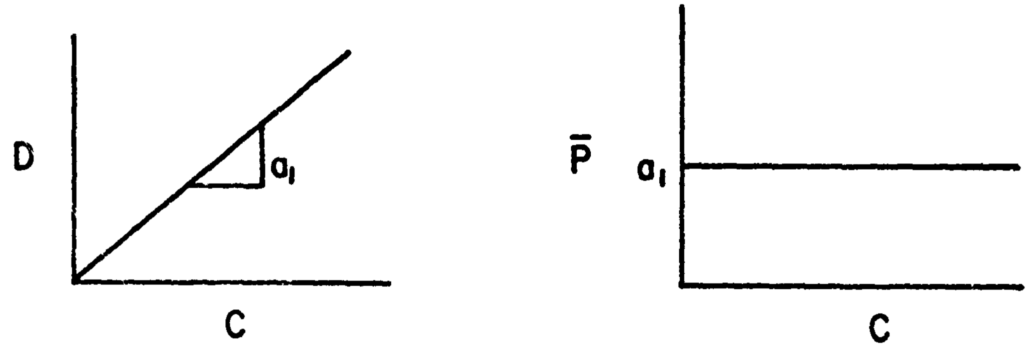


Figure B3. E&D costs/rates vs estimated cost of construction without economy of scale.

The basic model was generalized to reflect changes over time as follows:

Substituting Eq B1 into Eq B6

$$D_t = (1 + i_d)^t (a_0 + a_1 C)$$

Solving Eq B7 for C and substituting

$$Dt = (1 + i_d)^t \{a_0 + a_1 [\frac{1}{(1 + i_c)^t}] C_t\}$$

$$D_t = (1 + i_d)^t a_0 + [\frac{1 + i_d}{1 + i_c}]^t a_1 C_t \qquad \text{[Eq B8]}$$

It can be seen that the general model (Eq B8) is equivalent to the basic model (Eq B1) in the base year ($t = 0$).

Eq B8 was transformed into a linear equation using the approximation

$$(1 \pm \chi)^t \simeq \alpha_0 \pm \alpha_1 t$$

This approximation is quite good for small values of χ and when the maximum t is not large. The following nonlinear to linear transformations were made:

$$(1 \pm i_d)^t \simeq \alpha_0 \pm \alpha_1 t$$

$$\frac{1 + i_d}{1 + i_c}^t \simeq \beta_0 \pm \beta_1 t$$

Substituting the above linear approximations into Eq B8 gives

$$D_t = (\alpha_0 \pm \alpha_1 t) a_0 + (\beta_0 \pm \beta_1 t) a_1 C_t$$

$$D_t = \alpha_0 a_0 + \beta_0 a_1 C_t + \alpha_1 a_0 t + \beta_1 a_1 t C_t$$

$$D = b_0 + b_1 C + b_2 T + b_3 TC \qquad \text{[Eq B9]}$$

where $b_0 = \alpha_0 a_0$ $\quad b_2 = \alpha_1 a_0$
$b_1 = \beta_0 a_1$ $\quad b_3 = \beta_1 a_1$

and D and C are understood to be costs in the Tth year.

Eq B9 was then generalized for the ith Division/District to the complete model for E&D costs.

$$D_i = b_{0i} + b_{1i} C + b_{2i} T + b_{3i} TC \quad i = 0, 1, \ldots 8.$$

[Eq B10]

Model Analysis

Some basic goals were established before any computer regression runs were made:[1]

a. The final equation(s) should explain more than 90 percent of the variation ($R^2 > 0.90$)

b. The standard error of the estimate should be less than 10 percent of the mean (coefficient of variability < 10 percent)

c. All estimated coefficients should be statistically significant with $a = 0.05$

d. There should be no discernible pattern in the residuals.

The Statistical Package for the Social Sciences (SPSS)[2] was used to perform the regression analysis.

The first step in the analysis was to determine which, if any, of the Division/District models could

[1] N Draper and H. Smith, *Applied Regression Analysis* (John Wiley and Sons, Inc., 1966), p 238

[2] Norman H. Nie et al., *SPSS—Statistical Package for the Social Sciences* (McGraw Hill, 1975).

be combined either in part or wholly without affecting the statistical significance of the model. A modified "backward elimination" approach composed of the following steps was used:[3]

a. Regress the data against the model with all of the variables included

b. Use the F-statistic to determine whether the deletion of each variable would detract significantly from the model

c. Use Bartlett's Chi-square statistic to determine whether there is a significant difference between the residual variances of the districts

d. If residuals are homogeneous, the procedure is complete. If not, remove the data of the Division/District whose residuals are furthest from the residual mean variance and go to step a.

To accomplish the backward elimination, a dummy variable was used to indicate from which Division/District a particular data point came. Thus, the nine Divisions/Districts have a total of 36 possible variables, as shown in Eq B10.

The SPSS method required that a base District be chosen (and numbered zero); the coefficient for the other Districts represents their deviation from the base District. When the coefficients for all of the variables were determined through the regression analysis, the coefficients for any given district were added to the coefficients for the base District to estimate the total coefficient for the given District.

Confidence limits on forecasts are a function of the standard error of the estimate and the degrees of freedom related to that estimate. Degrees of freedom are in turn based on the number of observations (in this case, years of data). Each Division/District had only 10 observations. The backward elimination procedure was used to determine whether some Divisions/Districts behaved in generally the same way with respect to E&D costs. If so, their data could be pooled to increase the degrees of freedom, thus narrowing the distance between the prediction limits on any forecast.

The initial steps of the backward elimination procedure were performed as follows. There were initially 36 possible regression coefficients for the nine districts and their 90 data points, as shown in Eq B10. An enumeration tree (Figure B4) was established to explicitly show the removal of *a set* of regression coefficients. Each branch in the tree signifies the removal of the associated coefficient from all Districts *except* the base District. The basic premise for this is that if the data from the Districts are to be pooled, all Districts are performing in the same manner and thus the value of the coefficient for the base District (b_{jo}) adequately reflects that performance.

Table B2

Backward Elimination Results

Model ($i = 1, 8$)	R^2	df	F	α
1. All	.97247	54	—	—
2. Without b_{oi}	.96192	62	2.587	.02
3. Without b_{1i}	.96346	62	2.209	.04
4. Without b_{2i}	.96412	62	2.047	.06
5. Without b_{3i}	.96733	62	1.260	.28
6. Without b_{3o}, b_{3i}	.96589	63	1.434	.20

The regression analysis was initially performed with all 36 regression coefficients included to provide the best possible model fit. The value of R^2 was used to determine goodness of fit. Line 1 of Table B2 presents the results of this run. Next, runs of the model were made corresponding to each branch of the tree (Figure B5); i.e., one set of coefficients was removed at a time. Lines 2 through 5 of Table B2 show these results. The F-statistic was calculated as shown in

[3]Draper and Smith, p 167.

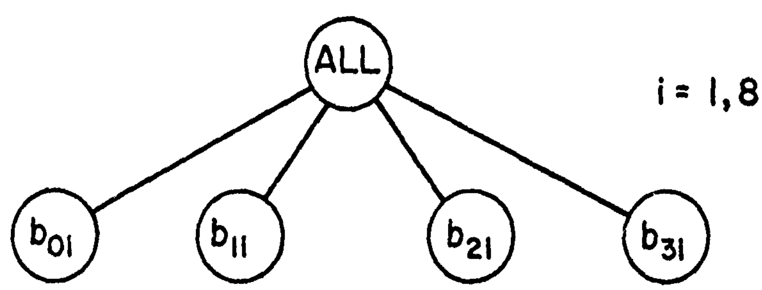


Figure B4. Removal tree for one variable at a time.

Eq B11 to determine whether there was a significant decrease in the model fit due to removal of the set of terms.

$$F = \frac{(R^2_{all} - R^2_{without})/(df_{without} - df_{all})}{(1 - R^2_{all})/df_{all}}$$

[Eq B11]

The value of α represents the risk incurred in rejecting the hypothesis that the removed terms are unnecessary, i.e., in concluding that the removed terms were necessary. A critical value of $\alpha = .10$ was used. According to the values obtained, only b_{3i}, the coefficient corresponding to the interaction term, can be removed without significantly reducing the model fit, ($\alpha = .28$).

The next step was to determine whether any additional terms could be removed. Since only the b_{3i} terms could be initially removed, b_{3o} is the only remaining possibility. Note that coefficients for the base District cannot be removed until the corresponding coefficients for all other districts have been removed. Figure B5 shows the removal of b_{3o}; the results are shown in Table B2, line 6. The value of $\alpha = .20$ indicates no significant reduction in model fit. For the data from the nine Districts under consideration, all of the interaction coefficients were the same and were assumed to be zero. The other nodes in Figure B5 have been fathomed (denoted by the underscoring) to indicate that since these terms could not be removed individually, they could obviously not be removed with other terms.

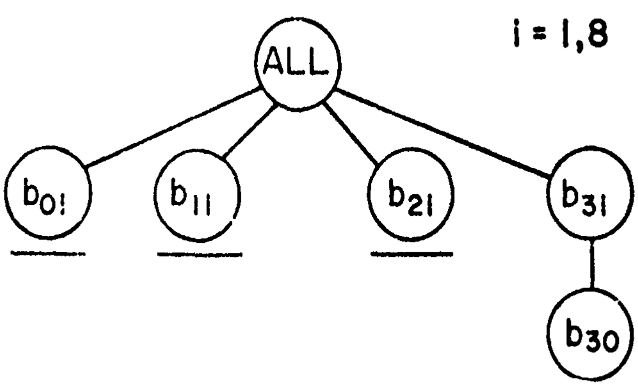


Figure B5. Removal tree for two variables at a time.

After the simplest model which would not significantly decrease the model fit for the nine Districts was determined, determining whether the data could, in fact, be pooled was the next step. This was done using Bartlett's Chi square[4] test to determine whether the residual variance differed from District to District. The variances of the residuals for each Division/District are shown in Table B3 (Step 1). According to Bartlett's test, the District-to-District variations differ significantly (probability = 1.0); therefore, the data from all nine Districts could not be pooled.

The data for the District/Division with the highest variation (in this case Pacific Ocean) were removed and the process was repeated.

The results of this process through five iterations are shown in Table B3. Four Districts had to be removed before the data for the remaining five could be pooled.

Table B3

Results of Test for Homogeneous Residual Variance Between Districts/Divisions

District/Division	Step 1	Step 2	Step 3	Step 4	Step 5
Baltimore	.01387	.01387	.04032	.01387	.01387
Fort Worth	.01744	.01744	.09565	.01744	.01744
New York	.02078	.01917	.02396	.02078	.02078
Omaha	.01705	.01705	.01811	.01705	.01705
Alaska	.01738	.01738	.01907	.01738	.01739
Sacramento	.03811	.03813	.03910	.03812	
Mobile	.18558	.18558	.18568		
Savannah	.19887	19887			
Pacific Ocean	1.93218				
Probability*	1.00	1.00	1.00	.30	.01

*Probability, based on Bartlett's Chi square statistic, that there is a significant difference between the variances from District to District

The same process was used to determine if any of the four removed Districts could be pooled. It was found that each of these had to be treated separately.

[4] G. W. Snedecor and W. G. Cochran, *Statistical Methods* (Iowa State University Press, 1971), pp 296-298